



**Barcelona
Supercomputing
Center**

Centro Nacional de Supercomputación

The Bone and The Marrow: Some Suggestions for the Future of Climate Prediction

Francisco Doblas-Reyes

R. Bernardello, R. Bilbao, A. Carreric, J. Cos, C. Delgado, M. Donat, N. Milders, B. Solaraju Murali, V. Torralba, E. Tourigny

28 March 2023



Funded by
the European Union



Outline: a collection of thoughts

Understanding the title:

- the bone and the marrow as a set of key elements for life
- chill to the marrow/to the bone

What we can do:

- with what we have
- with what climate change colleagues are doing
- with what other domains are doing

**Some of what else we can
do with what we have:
Transparency, standards,
guidance, user engagement**



**Barcelona
Supercomputing
Center**

Centro Nacional de Supercomputación

Non-trivial climatology definitions

ECMWF-Ext-Ens 2016

Period: 1996-2015 Month: April

Reference: ERA Int

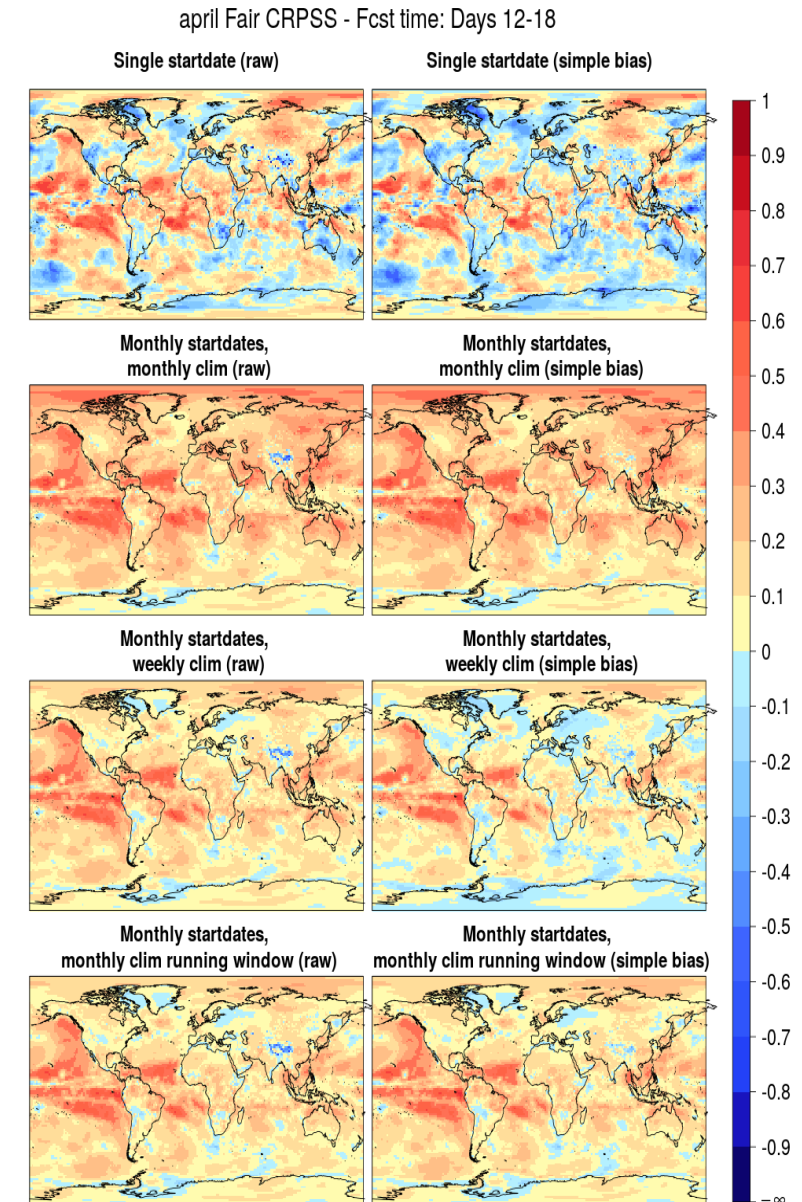
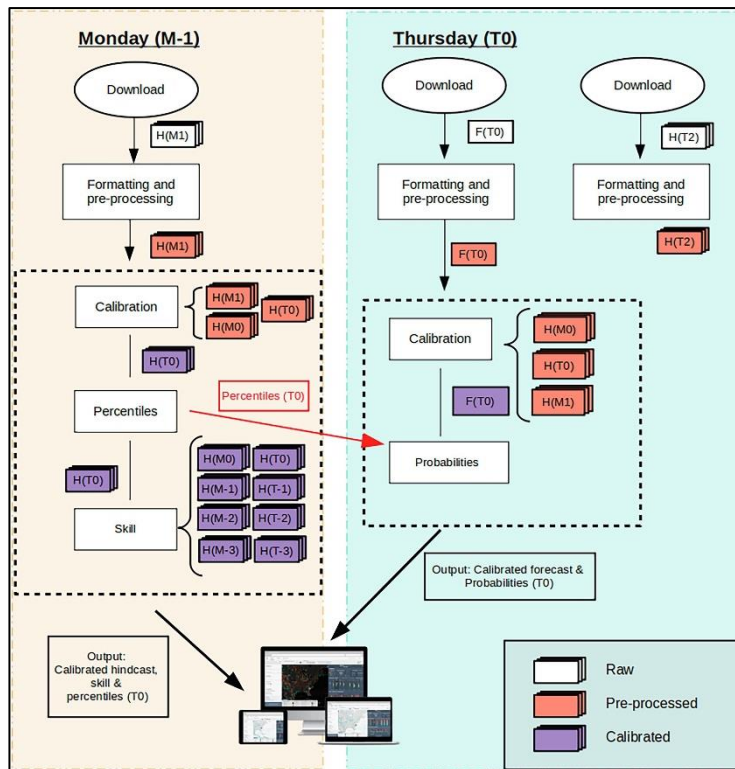
Variable: 2-m temperature

Weekly: too noisy

Monthly: good skill, but suspiciously high ...

Monthly running window, but weekly for the bias adjustment: lower skill, and too noisy for the adjustment

Monthly running window: more credible quality estimates



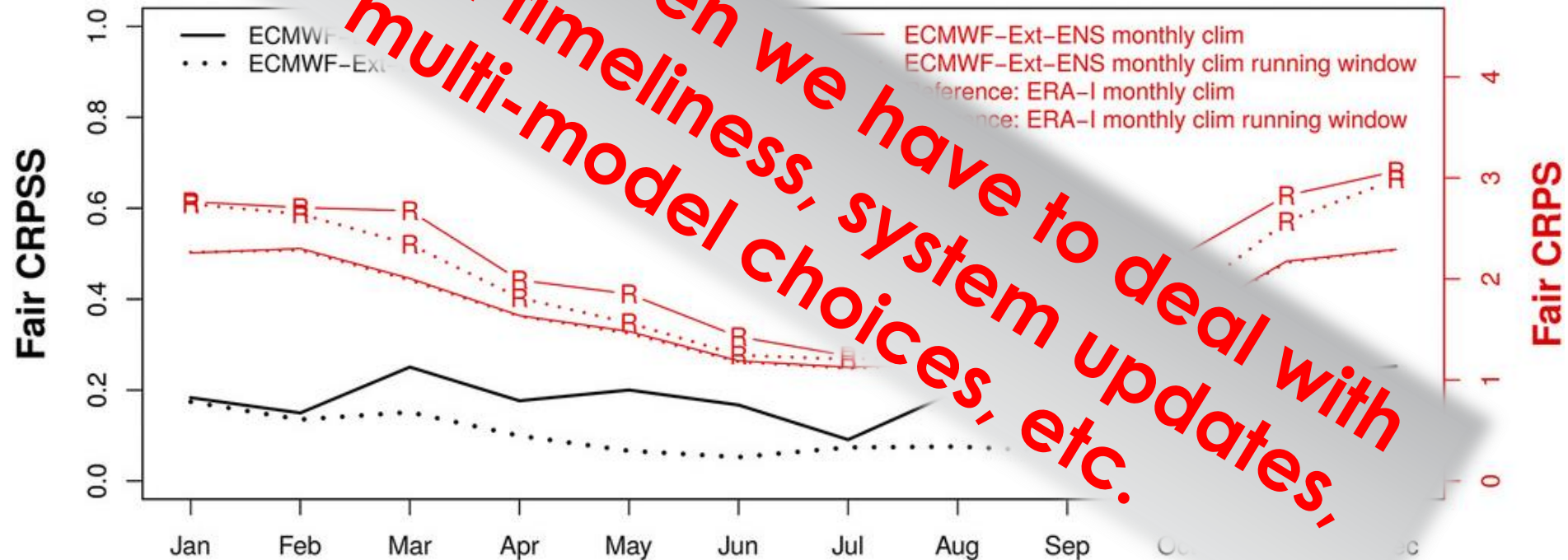
What for? Non-trivial climatology definitions

ECMWF-Ext-Ens 2016

Period: 1996-2015 Month

Reference: ERA Int

Variable: 2-m temper



Systematic forecast quality assessment

Near-Surface Air Temperature of CMCC System 35 (cross-validation = false, skill aggregation)

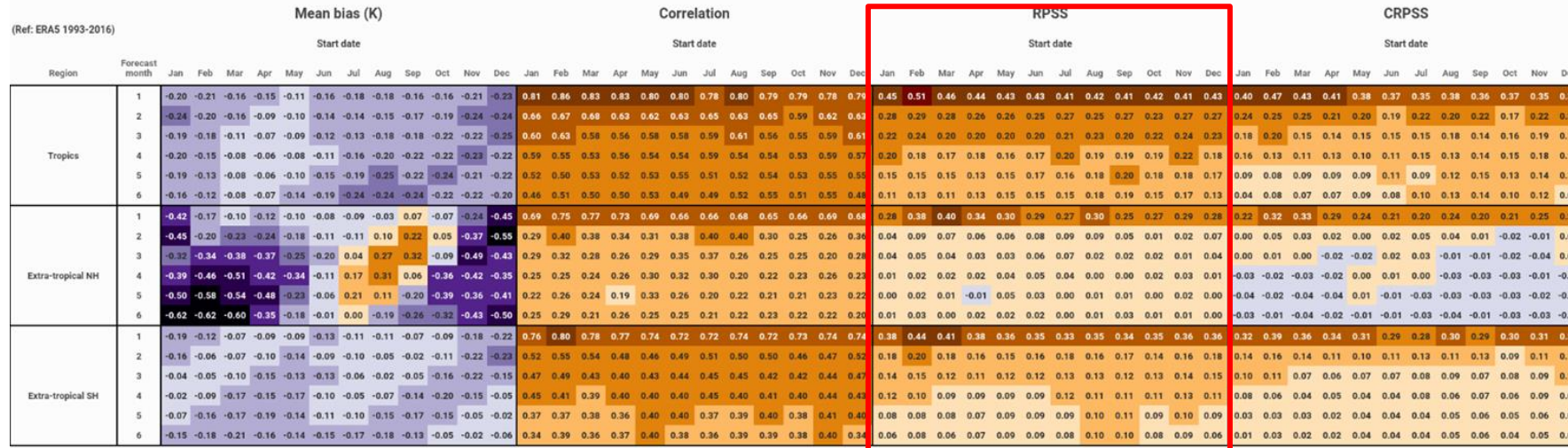
(Ref: ERA5 1993-2016)

		Mean bias (K)												Correlation												RPSS											
		Start date												Start date												Start date											
Region	Forecast month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Tropics	1	0.18	0.23	0.31	0.30	0.24	0.17	0.16	0.16	0.22	0.23	0.23	0.18	0.74	0.78	0.76	0.75	0.71	0.73	0.70	0.71	0.73	0.72	0.70	0.72	0.35	0.39	0.35	0.32	0.31	0.34	0.30	0.31	0.32	0.31	0.29	0.33
	2	0.18	0.29	0.30	0.26	0.13	0.11	0.11	0.17	0.20	0.18	0.15	0.15	0.61	0.62	0.63	0.59	0.57	0.59	0.59	0.58	0.60	0.56	0.57	0.60	0.23	0.23	0.22	0.21	0.20	0.19	0.20	0.20	0.22	0.19	0.22	0.23
	3	0.26	0.28	0.23	0.13	0.06	0.02	0.10	0.09	0.10	0.08	0.12	0.16	0.56	0.59	0.54	0.52	0.53	0.52	0.53	0.56	0.53	0.53	0.54	0.57	0.17	0.20	0.17	0.17	0.15	0.14	0.15	0.18	0.18	0.19	0.19	0.20
	4	0.27	0.20	0.10	0.04	-0.04	-0.02	-0.01	-0.02	0.00	0.06	0.15	0.25	0.55	0.51	0.49	0.50	0.48	0.48	0.52	0.50	0.52	0.53	0.55	0.52	0.16	0.15	0.12	0.14	0.10	0.11	0.15	0.15	0.18	0.19	0.19	0.15
	5	0.20	0.08	0.01	-0.07	-0.11	-0.14	-0.11	-0.11	0.00	0.09	0.24	0.28	0.48	0.47	0.47	0.45	0.46	0.47	0.47	0.49	0.51	0.52	0.50	0.52	0.11	0.11	0.11	0.09	0.09	0.11	0.13	0.15	0.17	0.18	0.15	0.15
	6	0.10	-0.01	-0.11	-0.15	-0.24	-0.23	-0.18	-0.11	0.02	0.18	0.27	0.22	0.41	0.45	0.44	0.42	0.46	0.43	0.48	0.49	0.52	0.49	0.50	0.45	0.07	0.09	0.08	0.07	0.10	0.10	0.14	0.16	0.17	0.14	0.14	0.10
Extra-tropical NH	1	-1.22	-1.15	-1.20	-1.12	-0.62	-0.21	0.23	0.15	-0.06	-0.20	-0.48	-1.03	0.61	0.66	0.67	0.63	0.61	0.61	0.57	0.59	0.56	0.57	0.60	0.62	0.21	0.28	0.29	0.24	0.24	0.24	0.19	0.20	0.18	0.18	0.21	0.22
	2	-0.79	-0.82	-1.05	-0.99	-0.79	-0.10	-0.01	-0.28	-0.25	-0.35	-0.71	-0.91	0.30	0.38	0.34	0.32	0.31	0.34	0.36	0.38	0.25	0.25	0.32	0.37	0.04	0.08	0.06	0.07	0.06	0.06	0.06	0.09	0.03	0.01	0.05	0.08
	3	-0.86	-1.11	-1.07	-1.06	-0.47	-0.19	-0.19	-0.28	-0.23	-0.59	-0.95	-0.86	0.29	0.32	0.28	0.27	0.28	0.35	0.34	0.27	0.19	0.24	0.24	0.29	0.04	0.05	0.05	0.05	0.04	0.06	0.06	0.04	0.00	0.02	0.04	0.05
	4	-1.27	-1.16	-1.24	-0.74	-0.52	-0.28	-0.04	-0.11	-0.51	-0.91	-0.92	-1.00	0.28	0.27	0.28	0.25	0.29	0.31	0.28	0.21	0.18	0.22	0.24	0.26	0.03	0.04	0.04	0.03	0.04	0.05	0.04	0.00	0.00	0.02	0.03	0.03
	5	-1.26	-1.31	-0.92	-0.78	-0.63	-0.07	0.15	-0.37	-0.90	-0.92	-1.08	-1.38	0.25	0.24	0.23	0.26	0.31	0.28	0.21	0.21	0.20	0.19	0.21	0.24	0.02	0.03	0.02	0.03	0.05	0.05	0.00	0.01	0.03	0.00	0.01	0.02
	6	-1.40	-1.05	-0.99	-0.88	-0.56	0.19	-0.20	-0.80	-0.90	-1.06	-1.48	-1.36	0.24	0.24	0.25	0.26	0.26	0.21	0.19	0.21	0.20	0.25	0.24	0.23	0.02	0.01	0.02	0.03	0.03	0.00	0.01	0.03	0.01	0.03	0.02	0.02
Extra-tropical SH	1	0.07	-0.04	-0.17	-0.24	-0.17	-0.10	0.03	0.00	-0.01	0.03	0.08	0.10	0.68	0.72	0.70	0.68	0.66	0.66	0.64	0.65	0.64	0.66	0.65	0.67	0.29	0.33	0.31	0.27	0.26	0.27	0.24	0.25	0.25	0.26	0.25	0.28
	2	0.07	-0.07	-0.22	-0.29	-0.31	-0.09	-0.07	-0.09	0.00	0.08	0.15	0.17	0.49	0.51	0.49	0.45	0.43	0.46	0.45	0.46	0.46	0.43	0.46	0.51	0.15	0.16	0.14	0.13	0.12	0.12	0.12	0.13	0.13	0.11	0.14	0.16
	3	-0.03	-0.19	-0.30	-0.40	-0.24	-0.19	-0.13	-0.08	0.05	0.13	0.12	0.06	0.45	0.45	0.41	0.38	0.40	0.41	0.41	0.43	0.39	0.41	0.43	0.45	0.11	0.12	0.10	0.09	0.09	0.08	0.09	0.11	0.10	0.11	0.12	0.13
	4	-0.13	-0.27	-0.44	-0.34	-0.35	-0.24	-0.10	0.01	0.10	0.10	0.03	-0.04	0.42	0.39	0.37	0.37	0.36	0.37	0.41	0.37	0.38	0.40	0.41	0.41	0.10	0.08	0.07	0.07	0.06	0.07	0.09	0.08	0.10	0.11	0.11	0.10
	5	-0.18	-0.40	-0.39	-0.44	-0.41	-0.19	0.01	0.07	0.06	0.00	-0.08	-0.12	0.35	0.35	0.35	0.34	0.35	0.38	0.36	0.36	0.38	0.38	0.38	0.39	0.06	0.06	0.06	0.05	0.05	0.07	0.07	0.08	0.10	0.09	0.08	0.08
	6	-0.30	-0.37	-0.50	-0.50	-0.39	-0.06	0.06	0.02	-0.04	-0.10	-0.16	-0.14	0.31	0.33	0.33	0.32	0.36	0.33	0.35	0.37	0.37	0.38	0.37	0.34	0.04	0.04	0.04	0.04	0.05	0.05	0.07	0.09	0.09	0.08	0.08	0.06



Systematic forecast quality assessment

Near-Surface Air Temperature of ECMWF System 5 (cross-validation = false, skill aggregation)



Near-Surface Air Temperature of ECMWF System 5 (cross-validation = true, skill aggregation)

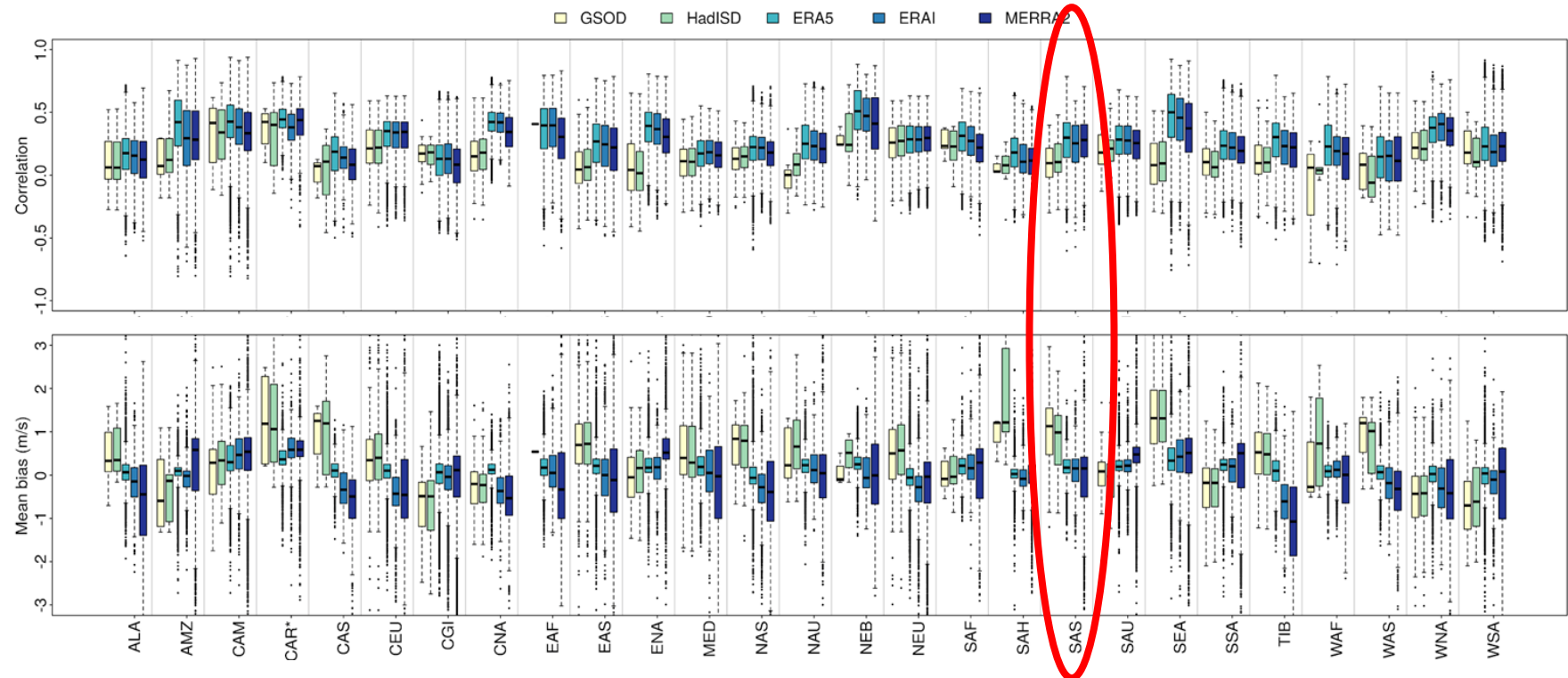
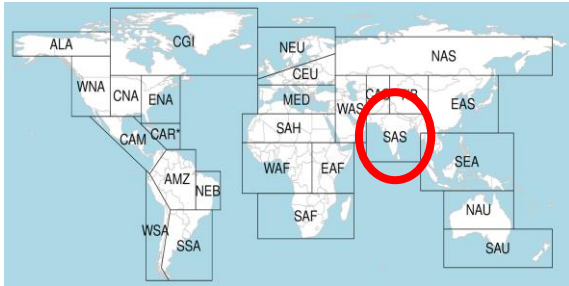


N. Milders (BSC)

Beware of the observational uncertainty

Verification with two ground-based observational datasets and three reanalyses. The use of both types of datasets is very informative for wind energy users as they use them for the development of impact models and the long-term resource assessments.

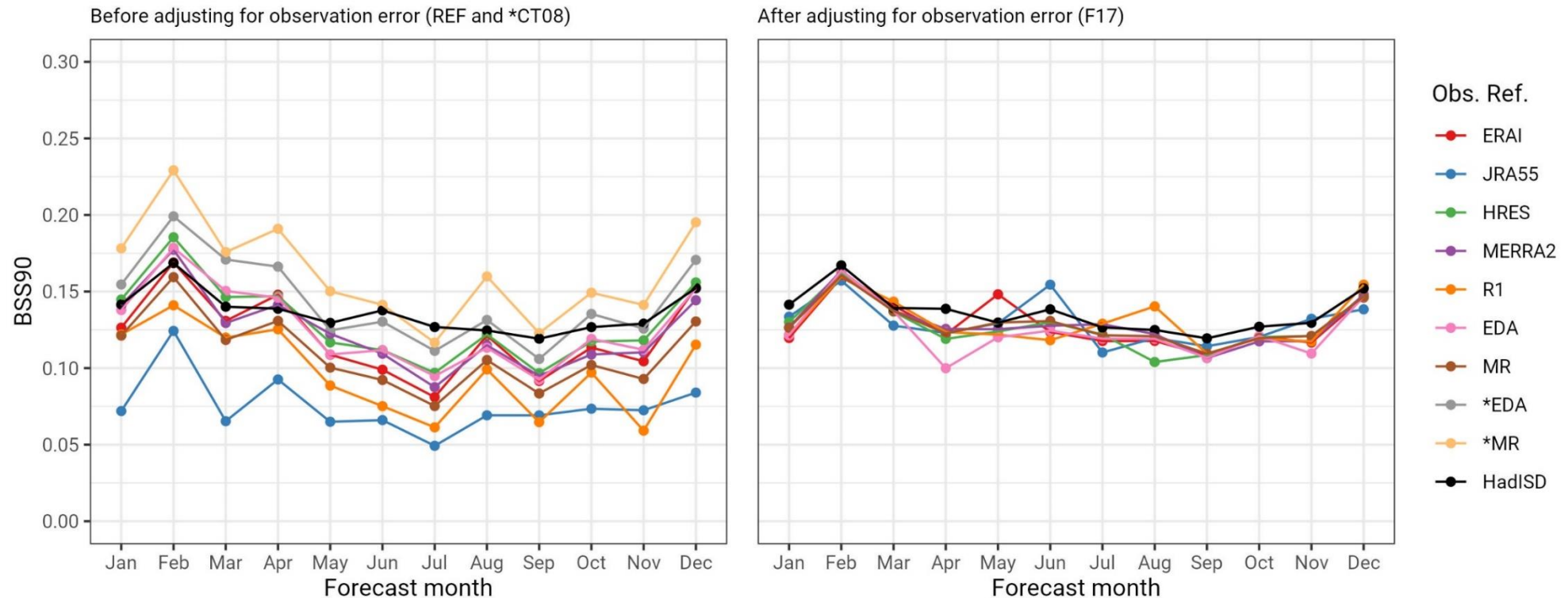
“Failing to account for the effects of observation error when deciding between two forecasting systems could lead to the wrong choice and a high opportunity cost” (Ferro, 2017)



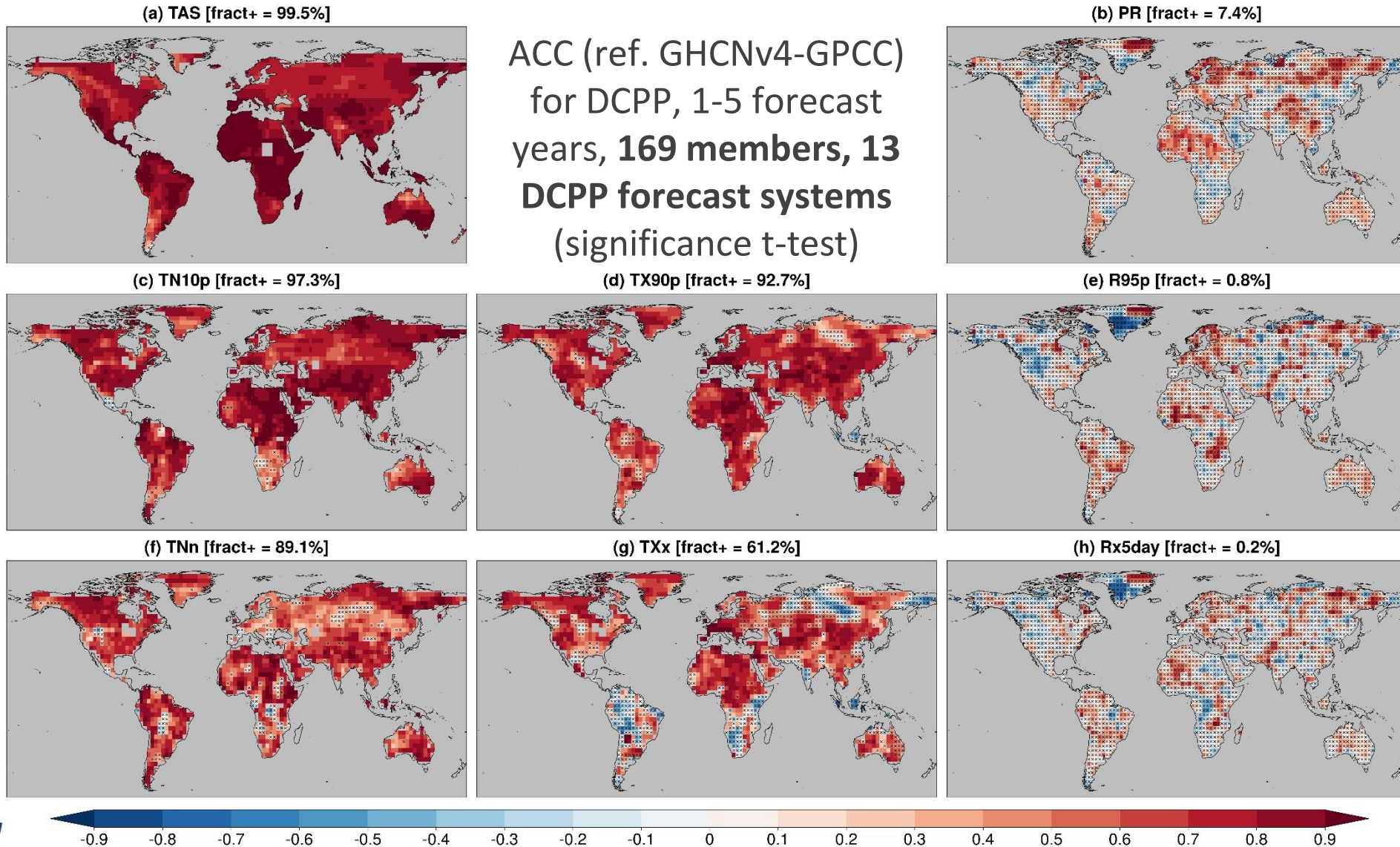
Observational uncertainty

Verification against several reanalysis-based references gives different results. Adjustment of the observational references to in-situ observations (HadISD) shows that this uncertainty can be reduced. 10 m wind speed, BSS90 for zero-month lead one-month forecasts from ECMWF S5 (1981-2017).

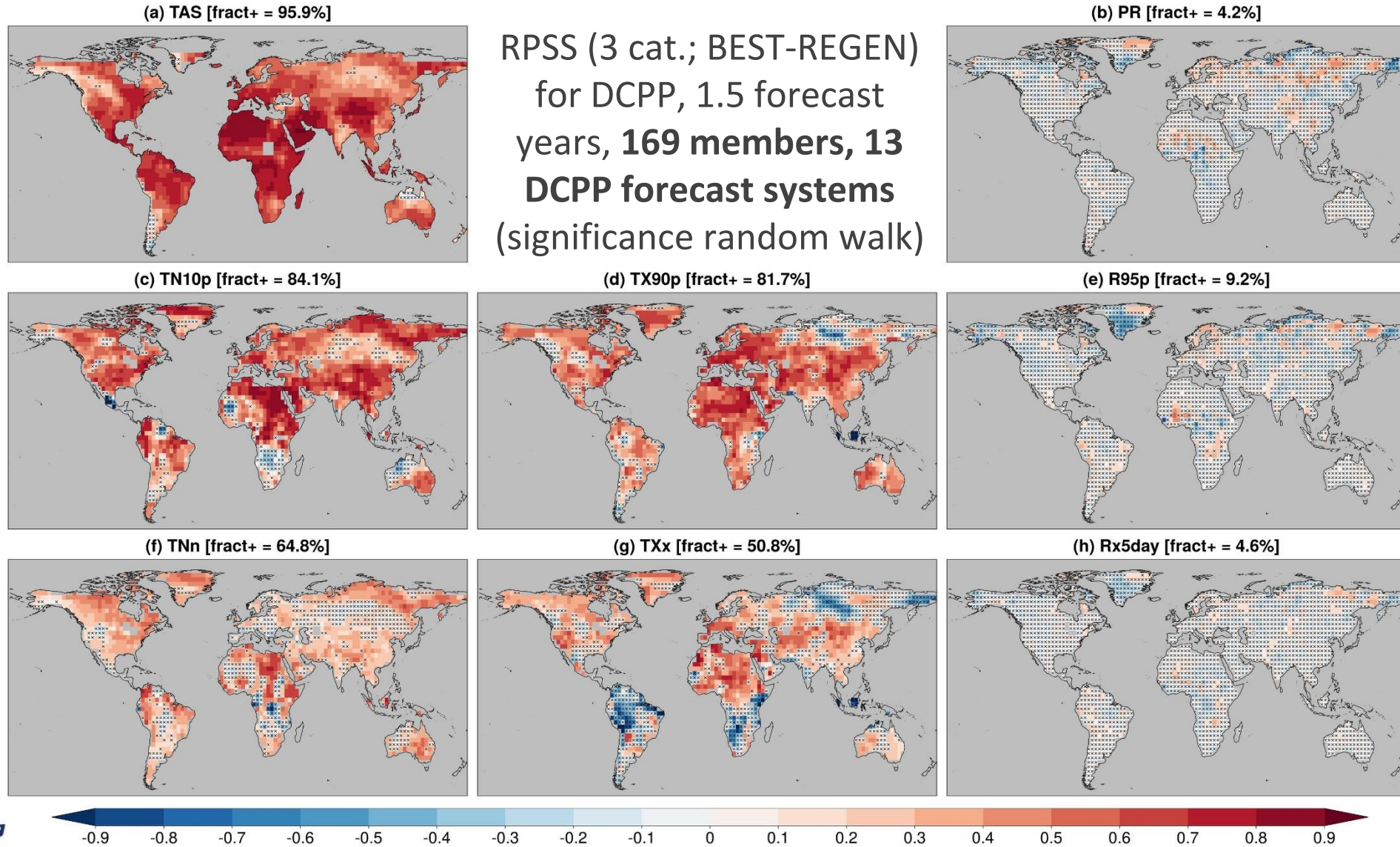
“Failing to account for the effects of observation error when deciding between two forecasting systems could lead to the wrong choice and a high opportunity cost” (Ferro, 2017)



Building multi-model systems



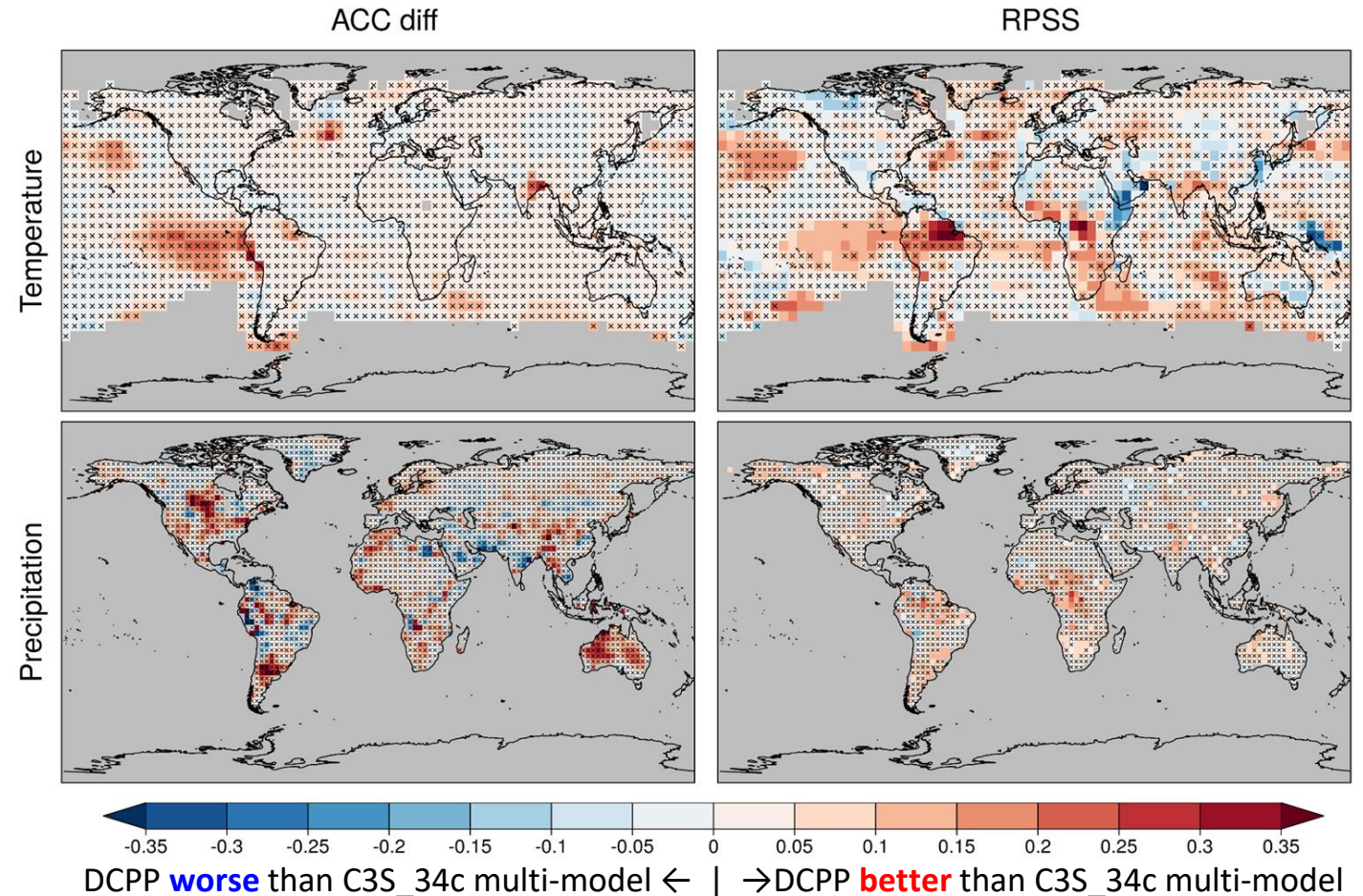
Building multi-model systems



Building multi-model systems

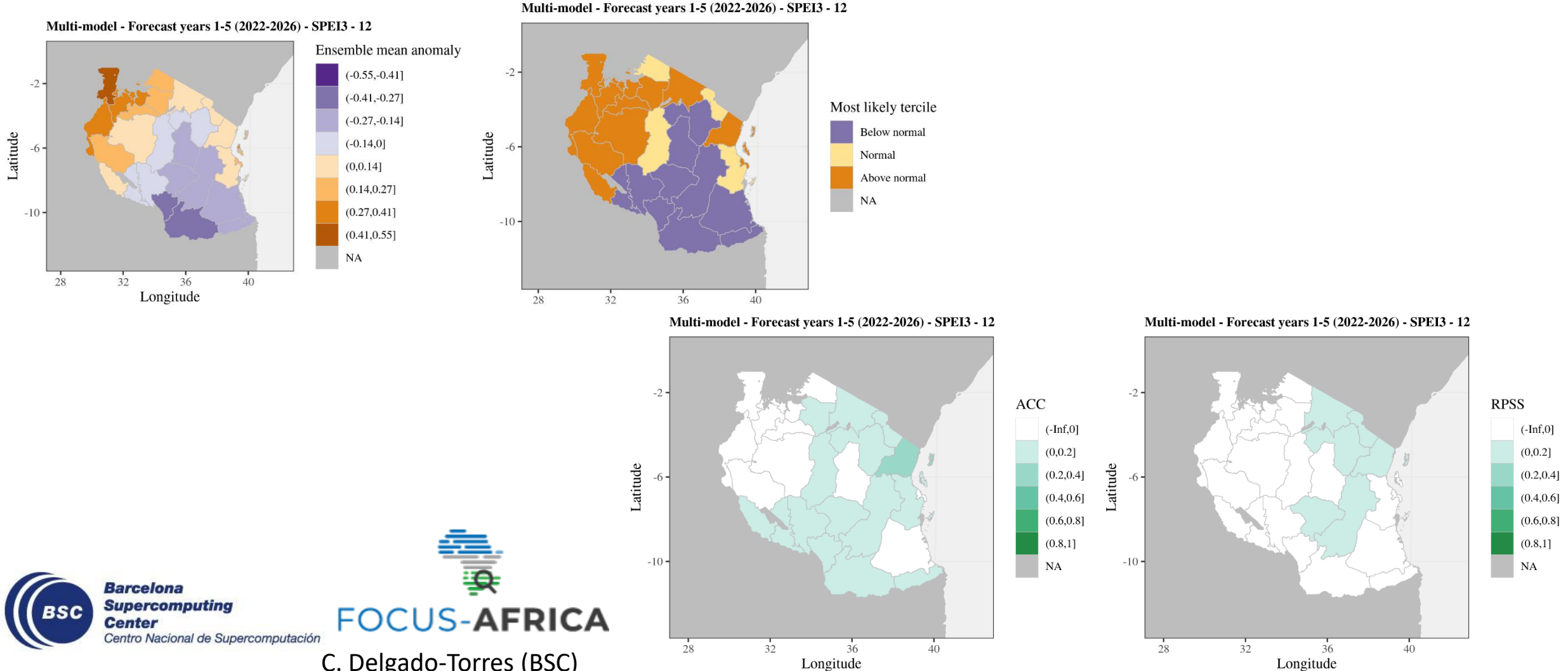
Systematic assessment of the multi-model decadal prediction forecast quality helps illustrating, among other things, the importance of a large enough operational multi-model.

Comparison between a research (DCPP, **169 members, 13 forecast systems**) and an operational (C3S_34c, **40 members, 4 forecast systems**, CMCC-CM2-SR5, EC-Earth3-i1, HadGEM3-GC3.1-MM and MPI-ESM1.2-HR).



Where to from here? User requirements

Downscaled SPEI3 (October to December) multi-model (three models) predictions for 2022-2026 over Tanzania to support agricultural decisions for maize. ERA5Land used as reference.



Lessons from research in user engagement

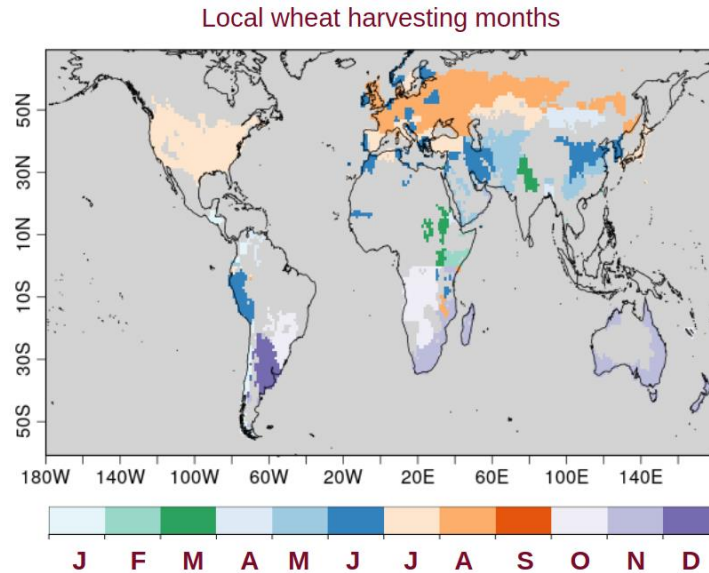
- Better links, collaboration and knowledge sharing among the increasing, and currently fragmented, community of practice. The current supply-driven approach is more likely to improve climate data, but not necessarily to contribute to better informed decisions. **Direct interaction would allow identifying improvement areas and emerging user needs** (role for climate services).
- Better interaction between **providers and users** and more effort to build local competences.
- Improving usefulness and usability of the products and **moving towards demand-driven** (instead of demand-relevant):
 - enhancing **credibility** with a more transparent approach of the corresponding climate services
 - enhancing **salience** by improving the skill, user interaction
 - enhancing **legitimacy** by creating independent quality assurance including the different stakeholders involved in the process
- Find a space for those **forecast intermediaries** (national and international) who have been using subjective methods, consider a balance between subjective and objective forecasts.



Early
Warnings
for All

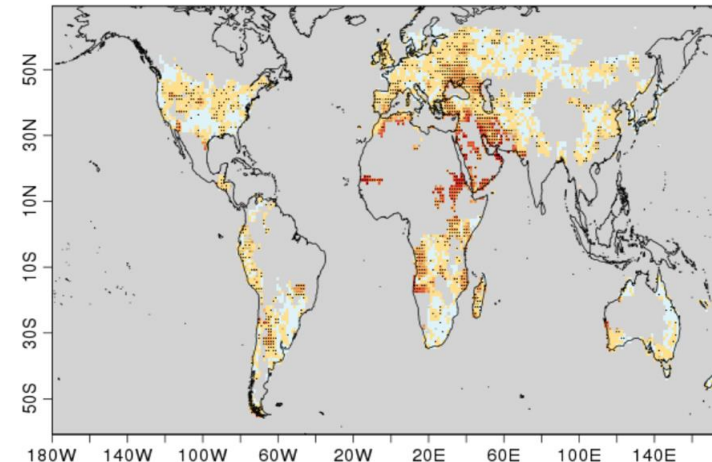
Forecasts for crop yield

WMO recognised global producing centres of decadal predictions contribute with the **definition of standards** for decadal predictions data and products, while C3S promotes the **evaluation** of the European multi-model and the illustration of the decadal prediction use in, among other sectors, the agricultural sector using indicators.

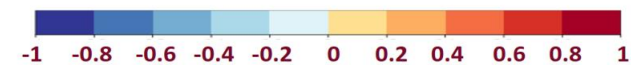


Indicators:

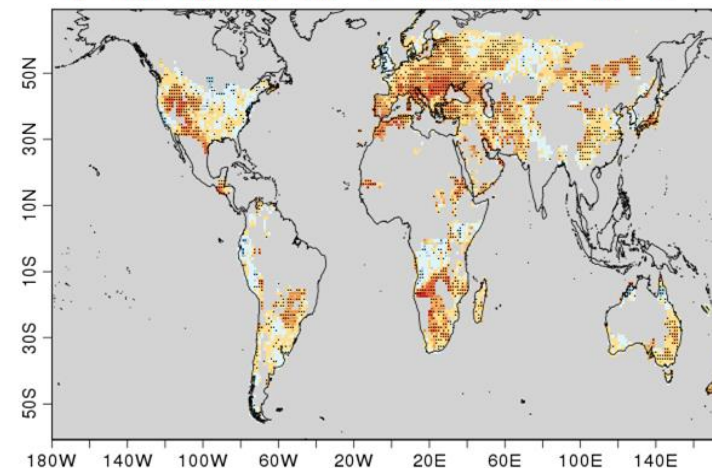
- Drought: Standardized Precipitation Evapotranspiration Index (**SPEI6**)
- Heat stress: Heat Magnitude Day Index (**HMDI3**)



SPEI6



RPSS (3 categories)



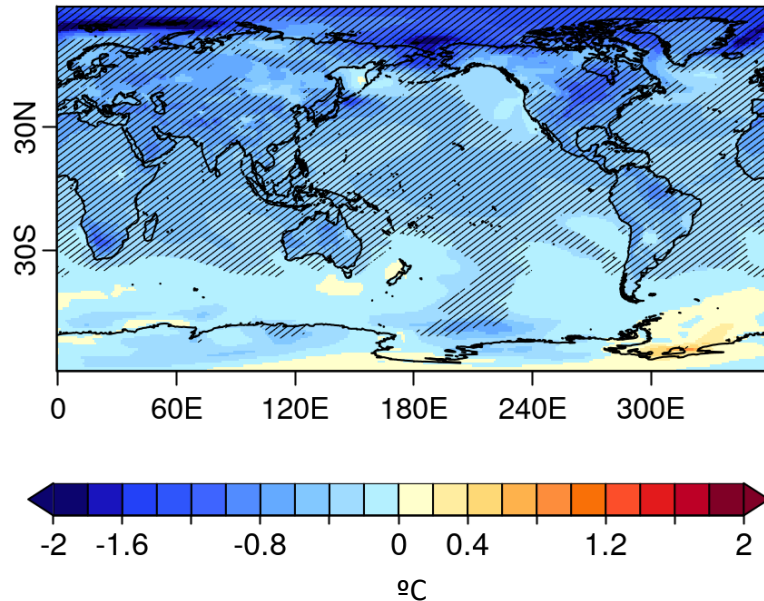
HMDI3

Volcanic influence

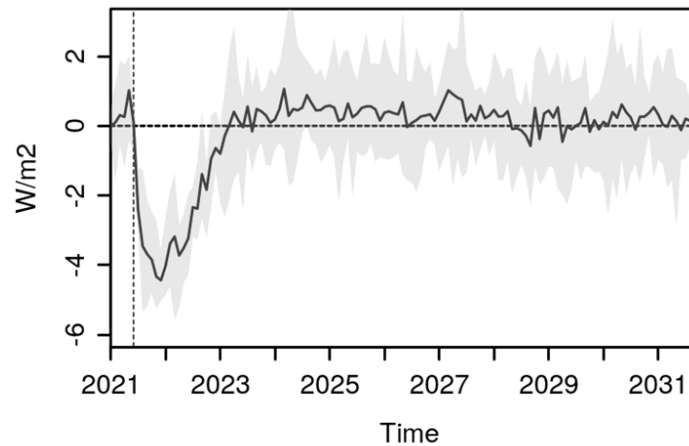
Results from the Decadal Prediction Volcanic Response Readiness Exercise (VolRes-RE).

A 2xEl Chichón eruption is set in April 2022 for the decadal forecast started in late 2021 and the difference with respect to DCP-P-A made.

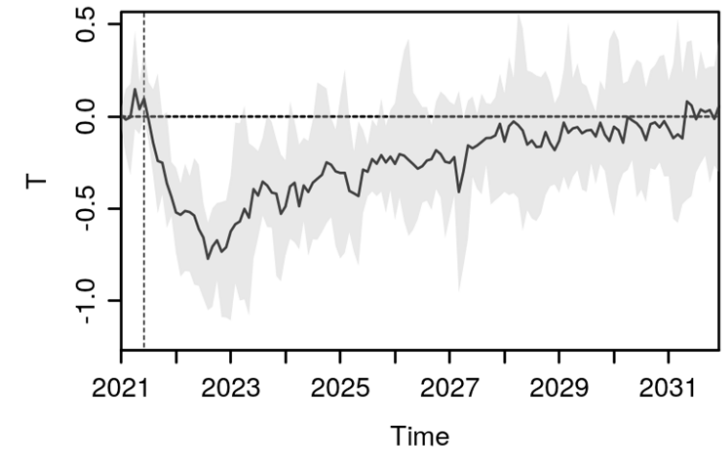
TAS YR2-5



a) Global Mean TOA



b) Global Mean Surface Temp.

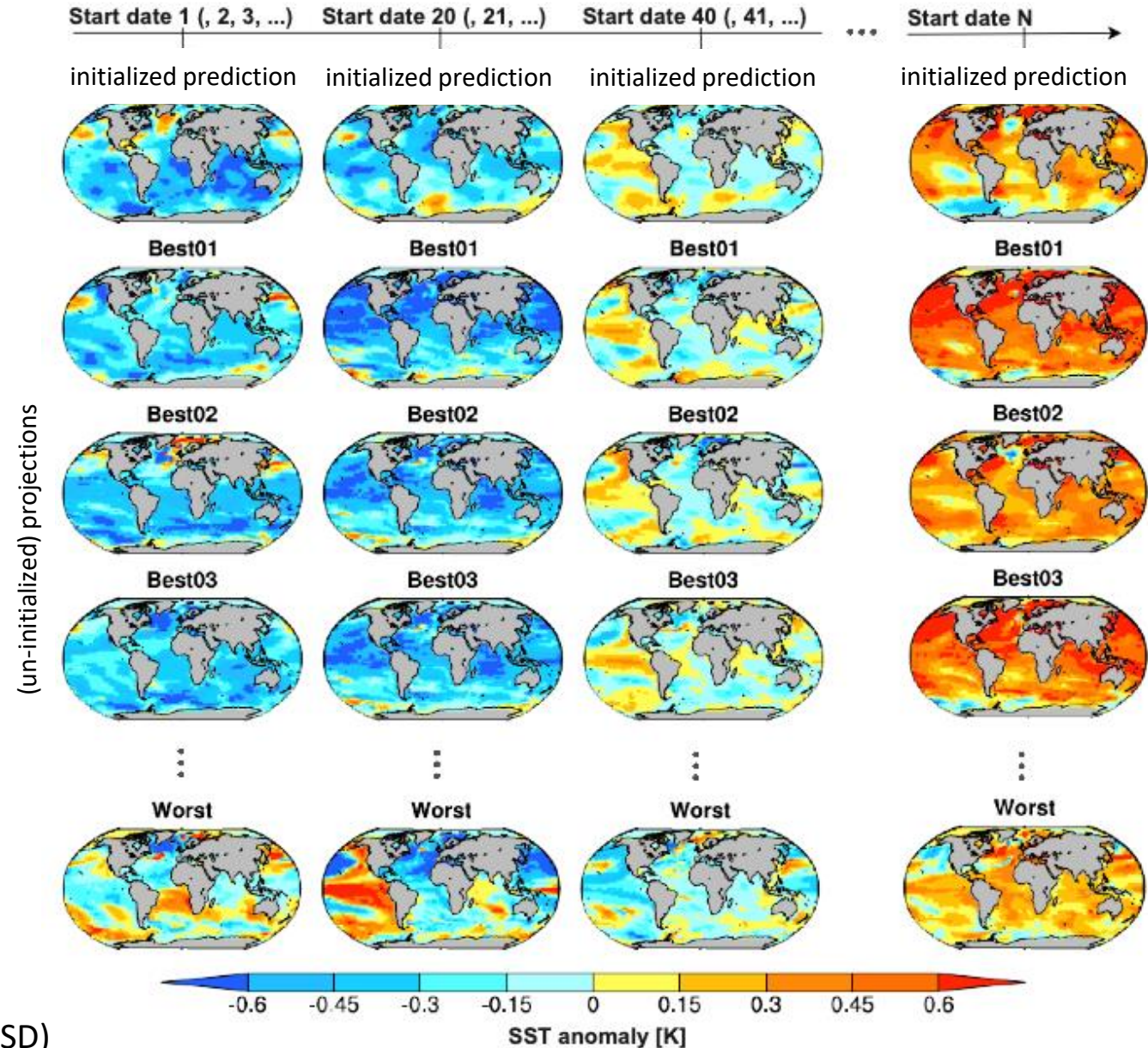
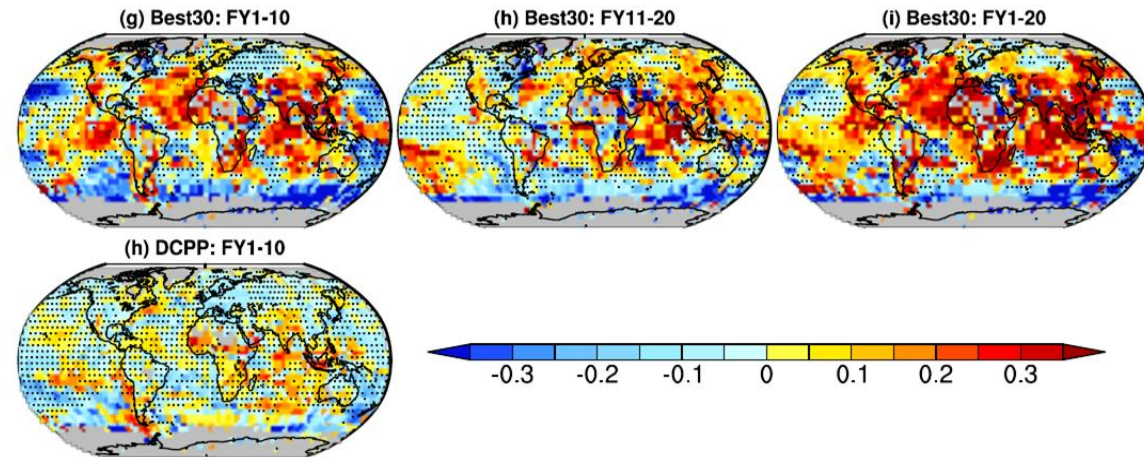


**Some of what we can do with
what happens around us:
Seamless information, more
realistic models, carbon cycle**

Near-term seamless climate information

Projections: **223 members** from 35 models
 Decadal predictions: **93 members** from 9 models
 Sub-selecting the 30 projections members in closest agreement with the predictions over years 1-9

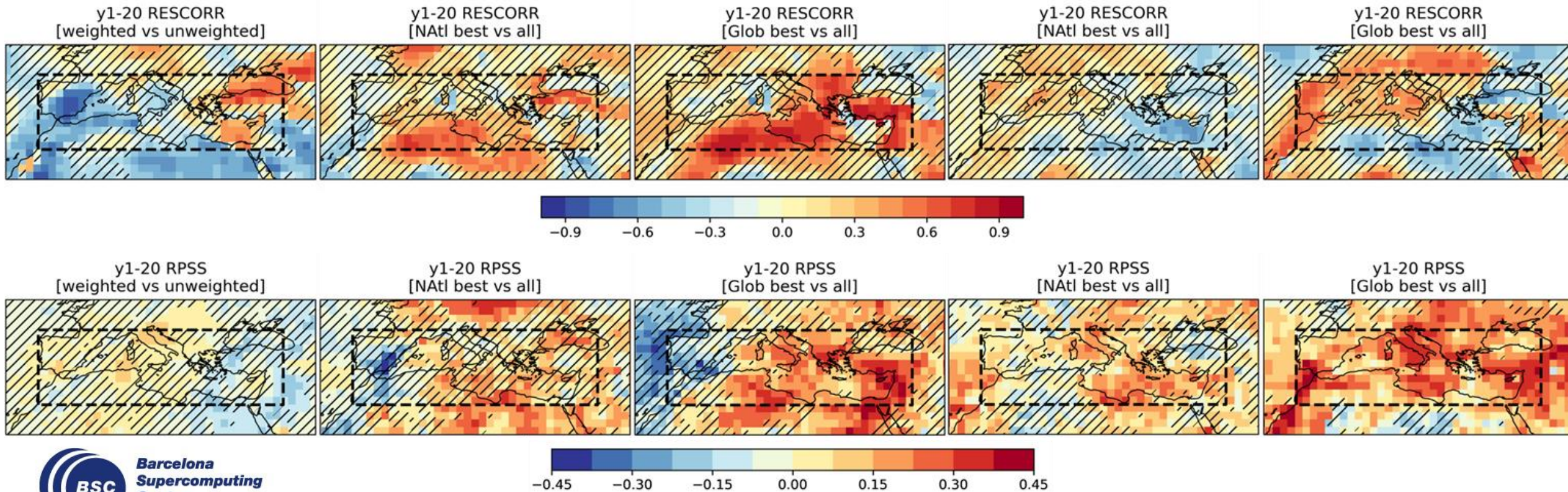
- There is substantial multi-decadal predictability by aligning internal variability
- Skill in constrained projections can be larger than in decadal predictions used to constrain



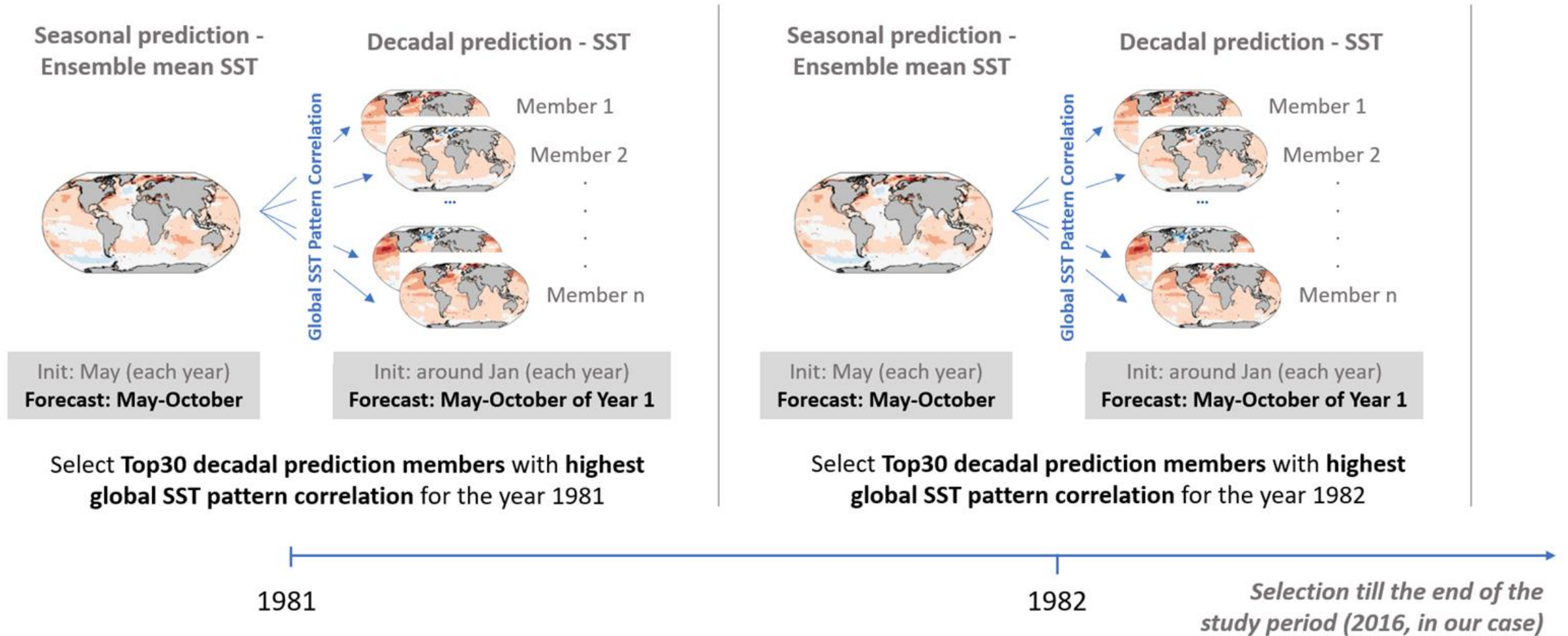
Near-term seamless climate information

Skill of 20-year projections using five constraining methods over the period 1970-2000

- Heterogeneity between methods
- It's fundamental to evaluate the constraining methods before claiming improvements in regional climate projections

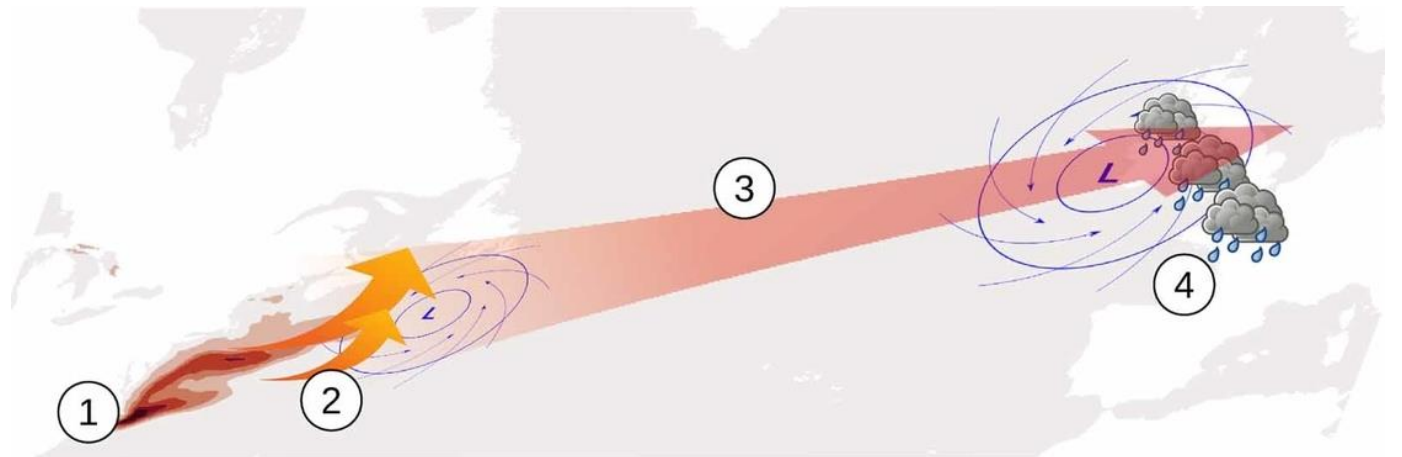
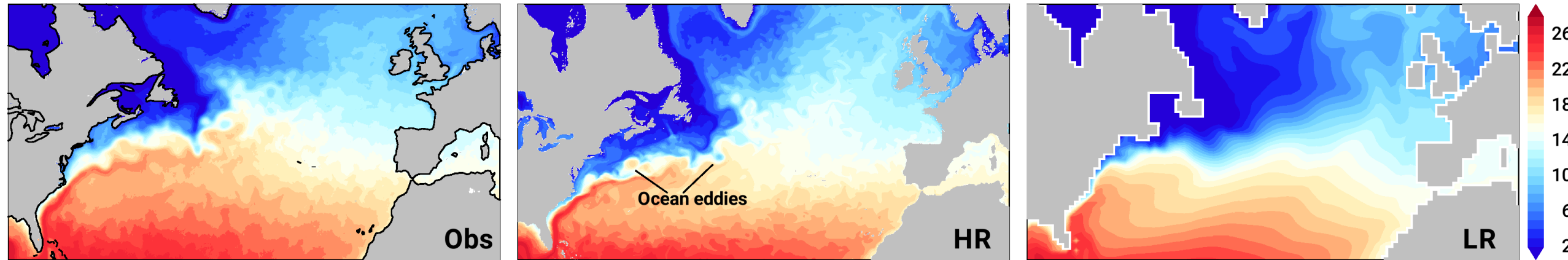


Merging seasonal and decadal predictions



New generation of climate models: resolution

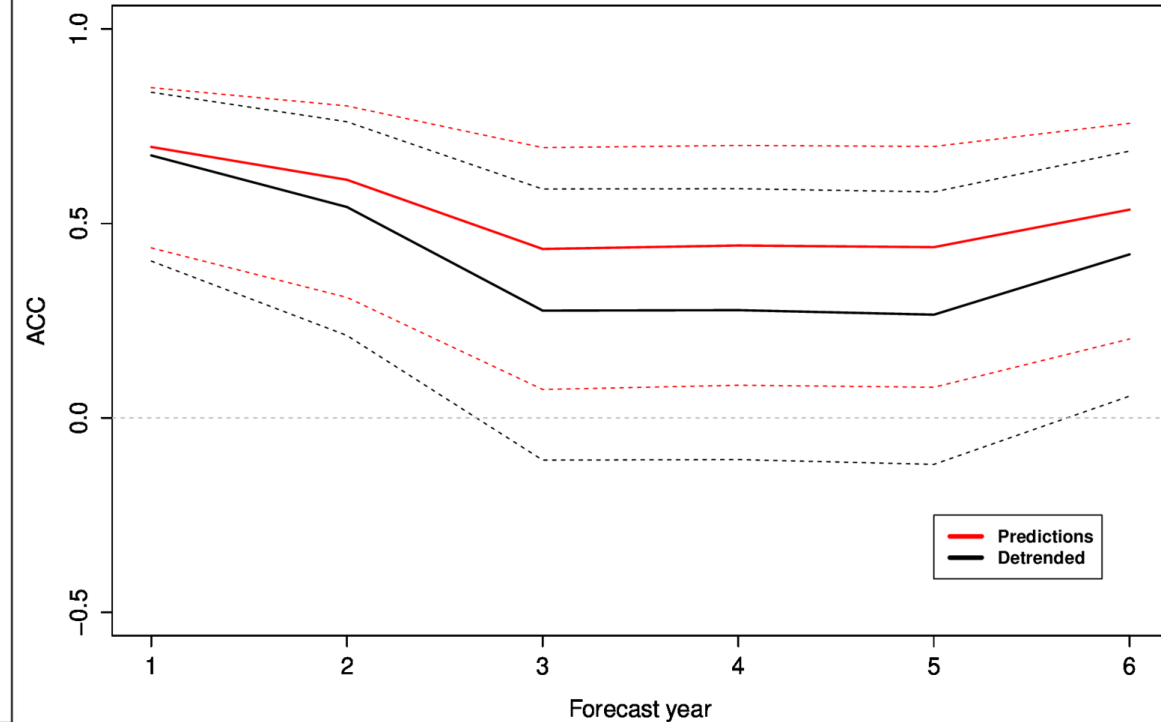
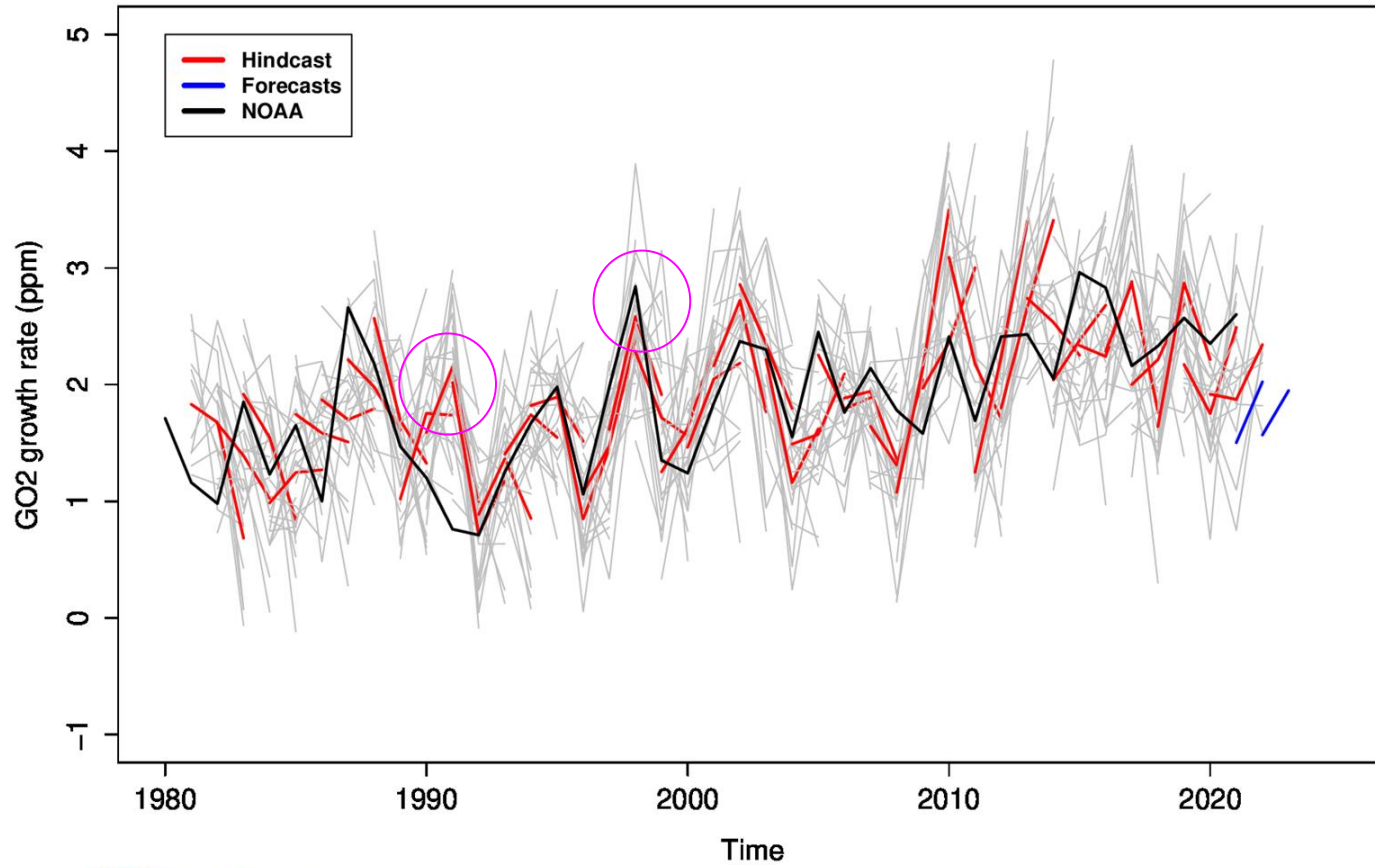
Eddy-rich, storm-resolving models (10-km or higher resolution) simulate a decrease in SST biases over the North Atlantic (right) and a northward shift in the Gulf Stream over the XXI Century that leads to rainfall increases in Europe (left), not found with traditional low-resolution models.



Carbon cycle predictions

Global atmospheric CO₂ growth rate predictions from EC-Earth3-CC. Initial conditions are obtained from off-line simulations.

There is skill up to two years. It then decreases mainly due to the land-atmosphere flux.



What gives the chill: Lack of interactivity, emergence of AI-based solutions



**Barcelona
Supercomputing
Center**

Centro Nacional de Supercomputación

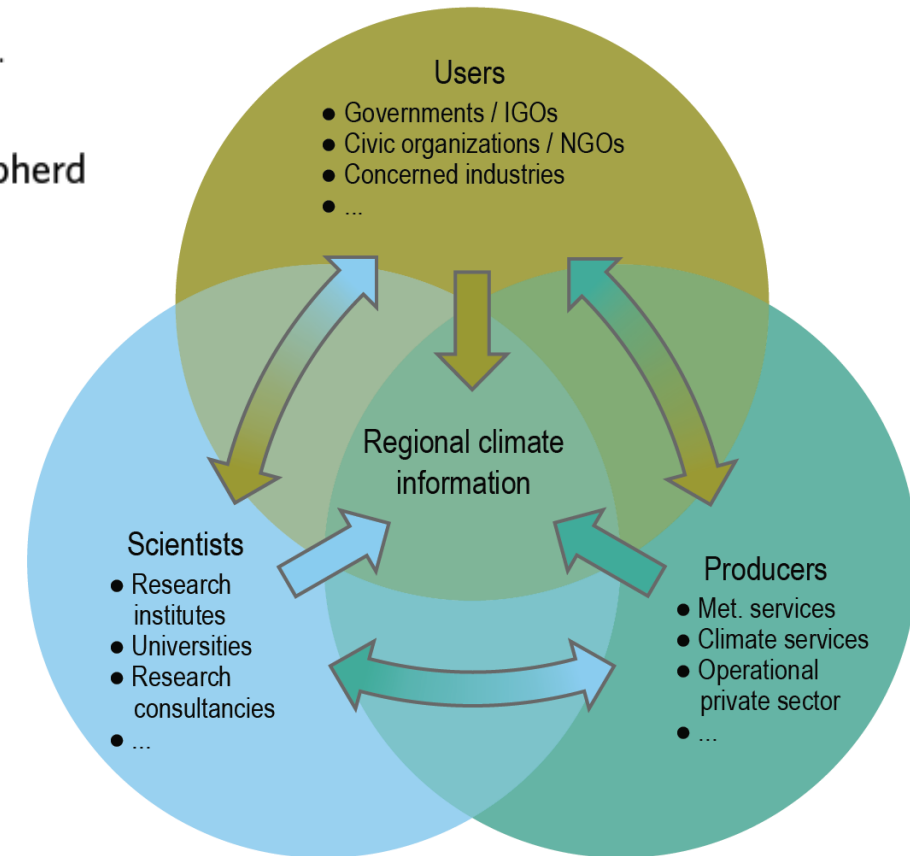
Actors and values in climate information

The value of values in climate science

To date, values are not widely acknowledged or discussed within physical climate science. Yet, effective management of values in physical climate science is required for the benefit of both science and society.

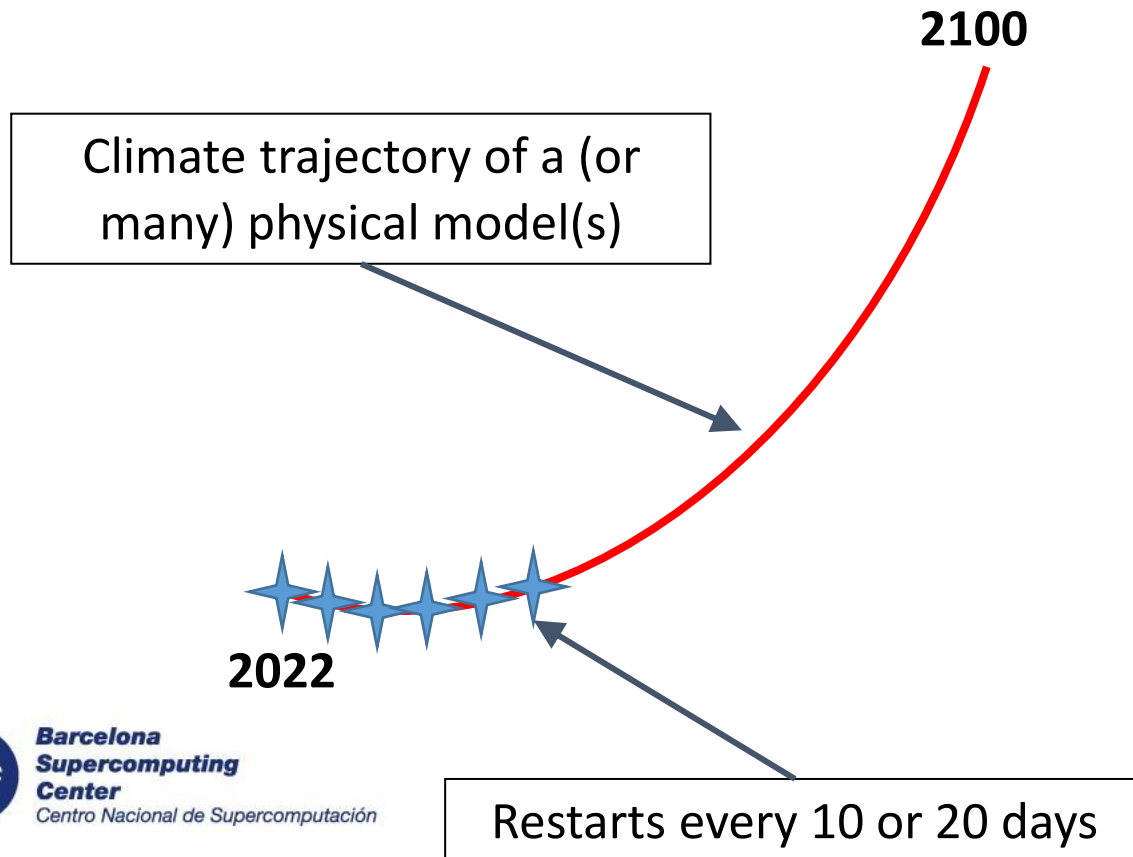
Karoliina Pulkkinen, Sabine Undorf, Frida Bender, Per Wikman-Svahn, Francisco Doblas-Reyes, Clare Flynn, Gabriele C. Hegerl, Aiden Jönsson, Gah-Kai Leung, Joe Roussos, Theodore G. Shepherd and Erica Thompson

This can only work if the user is offered true **interactivity** with the data sources.



Tethering: A necessary condition for interactivity

One of the ML-based weather forecast models (e.g., FourCastNet) could be trained on the trajectories of a (preferably several) process-based climate model (with high-frequency output). The tethering concept allows, using the climate model restarts, to recreate almost instantly (in inference mode as an interpolation between restarts) the model output requested by any user leading for true **interactivity**.



- The ML-based model can provide **downscaled** (super-resolution), **targeted** (user indicators), and **bias adjusted** data with large ensembles (if uncertainty handling is trustworthy).
- The physical model is a guarantee that there is a possibility to generate wildcards, i.e., unprecedented events plausible in a changing climate.

User interface to climate information

How long before a generative chatbot will be trained with the latest forecast information?

F

What is the probability of the occurrence of an El Niño event in October 2023?



As an AI language model, I do not have access to real-time weather data beyond my knowledge cut-off date of September 2021. The prediction for October 2023 is highly uncertain and depends on many complex changes in ocean temperatures and atmospheric patterns.

El Niño events typically occur every 2-7 years, but their frequency varies widely. However, based on historical records and current climate data, meteorologists and climate scientists may make predictions about the likelihood of an El Niño event occurring in October 2023.

F

What is the probability of lower than normal maize crop yield in Tanzania in the period 2023-2027?



The probability of lower than normal maize crop yield in Tanzania in the period 2023-2027 is difficult to predict with certainty as it depends on various factors, including weather conditions, pests and diseases, availability of inputs such as seeds and fertilizers, and socio-economic factors such as market prices and government policies.

However, there are several factors that could potentially affect maize crop yields in Tanzania. For instance, climate change could lead to unpredictable weather patterns, which could negatively impact crop yields. Additionally, pests and diseases could also affect crop yields if they are not effectively controlled.

GPT-4's capabilities and limitations create significant and novel safety challenges, and we believe careful study of these challenges is an important area of research given the potential societal impact. This report includes an extensive system card (after the Appendix) describing some of the risks we foresee around bias, disinformation, over-reliance, privacy, cybersecurity, proliferation, and more. It also describes interventions we made to mitigate potential harms from the deployment of GPT-4, including adversarial testing with domain experts, and a model-assisted safety pipeline.

Tanzania during the
and statistical
ging to accurately
ce them.
countries are

particularly vulnerable to crop failure and food insecurity, and efforts to support sustainable agricultural practices and increase resilience to climate change are essential to ensure food security and livelihoods.



**Barcelona
Supercomputing
Center**
Centro Nacional de Supercomputación

A collection of thoughts

Understanding the title:

- the bone and the marrow as a set of key elements for life
- chill to the marrow/to the bone

What we can do

- with what we have
- with what climate change colleagues are doing
- with what other domains are doing



**Barcelona
Supercomputing
Center**

Centro Nacional de Supercomputación